

Non-Gaussian dynamics from a simulation of a short peptide: Loop closure rates and effective diffusion coefficients

John J. Portman

Center for Nonlinear Studies and Theoretical Division,
Los Alamos National Laboratory, Los Alamos, New Mexico 87501
(Dated: October 31, 2002; laur-02-4552)

Intrachain contact formation rates, fundamental to the dynamics of biopolymer self-organization such as protein folding, can be monitored in the laboratory through fluorescence quenching measurements. The common approximations for the intrachain contact rate given by the theory of Szabo, Schulten, and Schulten (SSS) [J. Chem. Phys. **72**, 4350 (1980)] and Wilemski-Fixman (WF) [J. Chem. Phys. **60**, 878 (1973)] are shown to be complementary variational bounds: the SSS and WF approximations are lower and upper bounds, respectively, on the mean first contact times. As reported in the literature, the SSS approximation requires an effective diffusion coefficient 10 to 100 times smaller than expected to fit experimentally measured quenching rates. An all atom molecular dynamics simulation of an eleven residue peptide sequence in explicit water is analyzed to investigate the source of this surprising parameter value. The simulated diffusion limited contact time is approximately 6 ns for a reaction radius of 4 Å for solvent viscosity corresponding to that of water at 298 K and 1 atm ($\eta = 1.0$ cP). In analytical work, the polymer is typically modeled by a Gaussian chain of effective monomers. Compared to Gaussian dynamics, the simulated end-to-end distance autocorrelation has a much slower relaxation. The long time behavior of the distance autocorrelation function can be approximated by a Gaussian model in which the monomer diffusion coefficient D_0 is reduced to $D_0/6$. This value of the diffusion coefficient brings the mean end-to-end contact time from analytical approximations and simulation into agreement in the sense that the SSS and WF approximations bracket the simulated mean first contact time. Attention to the non-Gaussian nature of the dynamics has direct implications for the development of improved analytical models.

Introduction

Intrachain contact formation is fundamental to our understanding of biopolymer dynamics involved in molecular self-organization. Motivated primarily by the relevance to protein folding, several groups have recently measured the quenching rates of the terminal residues on short peptides using fluorescent probes with short range quenching.^{1,2,3} In these experiments, the fluorescence decay rate characterizes the rate of contact formation in peptides for which the dynamics should be less complicated (and hence, easier to interpret) than the folding dynamics of natural protein sequences. Lapidus *et al.*¹ use a simple polymer model with a single dynamical variable to understand the measured rates, finding that the effective diffusion constant needed to fit the data is approximately 30 times smaller than the expected relative coefficient for free diffusion of the two terminal residues (assuming this to be $1.5 \times 10^{-5} \text{ cm}^2/\text{s}$). A small effective diffusion coefficient appears to be a rather general result, consistent over different molecules and types of probes. Many years ago, Haas *et al.*⁴ also reported surprisingly small diffusion coefficients using similar polymer models (see also Ref. 5), but here the quenching mechanism of the terminal monomers of the peptides is through the longer range fluorescence resonance energy transfer (FRET). Recently, Wallace *et al.*⁶ applied this analysis to FRET quenching experiments in DNA, concluding that the effective monomer diffusion coefficient was 1000 times slower than reported for peptides.

Interpretation of this small effective diffusion coefficient is particularly relevant to modern protein folding research. The energy landscape theory of protein folding⁷ along with advances in experimental techniques that probe faster timescales has encouraged some researchers to focus on proteins with the simplest kinetics, the so-called fast folding proteins.^{8,9} There has been considerable progress in understanding these fast folding proteins from both experimental and theoretical approaches. For this group of proteins, the transition state ensemble probed experimentally through mutational studies pioneered by Fersht¹⁰ can be modeled fairly accurately by simple polymer based models with interactions between distant residues specified by the native state topology.^{11,12,13,14,15} Success of modeling the barrier crossing dynamics that determine folding rates is far less clear; the few theories of the folding dynamics proposed so far^{16,17,18,19} could greatly benefit from more experimentally supported microscopic parameters characterizing the peptide dynamics. Such fundamental timescales and model parameters can be obtained in principle from measured intrachain quenching rates and more accurate theoretical modeling.

Effective parameters are relevant to specific models. A common model for polymer dynamics approximates the polymer backbone connectivity by harmonic bonds between effective monomers. As emphasized by Zwanzig,²⁰ a general harmonic potential energy can be defined through the static correlations between the monomers. Interactions between distant monomers do not enter explicitly into the model except possibly (and indirectly)

through the static correlations defining the Gaussian potential. The dynamics of these phantom chain models obey Gaussian statistics. In particular, the dynamics of

the end-to-end vector $\mathbf{R}(t)$ is completely described by the Gaussian Green's function (assuming isotropy)

$$G(\mathbf{R}|\mathbf{R}_0) = \left(\frac{3}{2\pi \langle R^2 \rangle [1 - \phi^2(t)]} \right)^{3/2} \exp \left[-\frac{3}{2 \langle R^2 \rangle} \frac{|\mathbf{R} - \phi(t)\mathbf{R}_0|^2}{1 - \phi^2(t)} \right]. \quad (1)$$

Here, $\phi(t)$ is the equilibrium vector correlation

$$\phi(t) = \frac{\langle \mathbf{R}(t) \cdot \mathbf{R}(0) \rangle}{\langle R^2 \rangle}, \quad (2)$$

and $\langle R^2 \rangle$ is the equilibrium mean square end-to-end distance that characterizes the size of the polymer through the Gaussian equilibrium distribution

$$P_{\text{eq}}(\mathbf{R}) = \left(\frac{3}{2\pi \langle R^2 \rangle} \right)^{3/2} \exp[-\beta U(\mathbf{R})], \quad \beta U(\mathbf{R}) = \frac{3R^2}{2 \langle R^2 \rangle}. \quad (3)$$

As long as the Gaussian nature of the dynamics is retained, the only relevant difference between different Gaussian polymer models is through the static size $\langle R^2 \rangle$ and the equilibrium pair correlation $\phi(t)$. Neglecting friction with memory,^{20,21} $\phi(t)$ is typically expressed as a sum of exponentials with a wide distribution of decay times. Another common approximation treats $\mathbf{R}(t)$ as the only dynamical variable, replacing $\phi(t)$ in Eq.(1) with a single exponential. This single variable approximation to the multi-bead dynamics is the one that has been used to analyze the fluorescence quenching experiments.

Some explanations for the small effective diffusion coefficient required to fit measured quenching rates can be incorporated into Gaussian models. For example, Haas *et al.*⁴ suggested that the reduced effective diffusion constant may be evidence for internal friction of the chain (independent of the solvent viscosity). Internal friction designed to capture the timescale of dihedral angle basin hopping has been introduced through a modified friction matrix while retaining the Gaussian description of the dynamics.^{22,23} Alternatively, Lapidus *et al.*¹ noted that the reduced effective diffusion constant is reminiscent of diffusion in a one-dimensional rugged potential well as discussed by Zwanzig.²⁴ Within the single variable approximation, averaging over the rugged energy potential can lead to dynamics described by diffusion on a smooth potential with a reduced effective diffusion constant. Both local chain dynamics or energetic ruggedness can provide additional friction that effectively slows the dynamics of the Gaussian chain. A main focus of this paper is to determine whether a Gaussian chain with modified parameters can capture timescale of the polymer dynamics suggested by the small effective diffusion coefficient.

There are certainly other mechanisms that could slow the dynamics, some even leading to glassy behavior. The ruggedness in Zwanzig's one dimensional model are local extrema in the potential along the reaction coordinate, sometimes referred to as "longitudinal ruggedness." Alternatively, the Bryngelson and Wolynes theory of protein folding dynamics²⁵ accounts for "transverse ruggedness" (the coupling to hidden or additional degrees of freedom) that may induce kinetic traps along the reaction coordinate. Interactions between monomers along the polymer chain can also generate a form of friction. Mode-coupling theories predict that the dynamics of random heteropolymers may become non-ergodic,^{26,27} similar to spin glass dynamics below the dynamical glass transition temperature.²⁸ Although one would not expect a short peptide to exhibit glassy behavior, similar mechanisms may still be relevant to some degree in this case as well.

The general problem of intrachain quenching is a difficult one that has not been solved explicitly even for a Gaussian chain.²⁹ Two widely used approximations are the closure approximation proposed by Wilemski and Fixman³¹ (WF) and the single variable approximation solved by Szabo, Schulten, and Schulten³² (SSS). In this paper, it is shown that the diffusion limited reaction rate from these two approximations are related through a variational expression: the WF and SSS approximations are upper and lower bounds, respectively, to the mean first passage time of end-to-end contact of the original chain. Since the single variable approximation used to interpret the experimental measurements is a lower bound, it is possible that the SSS approximation requires a small effective diffusion constant to fit the experimental quenching time because it is a poor approximation to the exact mean contact time of the multi-bead chain. Similarly, the WF approximation (had it been used) requires a larger diffusion constant to fit the data, since it gives a longer contact time than the SSS approximation.

The accuracy of these approximations have been assessed by the comparison to simulations of a Rouse chain (monomers sequentially connected by harmonic springs).^{33,34,35} The results reported by Pastor *et al.*³³ show that the SSS approximation underestimates the Rouse chain mean contact time by roughly a factor of three (although this factor is just an example as it depends strongly on chain length and contact distance).

Consequently, the accuracy of the lower bound may at least account for part of the small effective diffusion constant that needs clarification.

In order to develop improved models informed by measured quenching rates, it is crucial to clarify the way in which the effective diffusion coefficient reflects the timescale of the polymer dynamics. This question is addressed in this paper through an analysis molecular dynamics simulation of a realistic peptide. To connect with analytic polymer models, a general Gaussian chain is chosen to define the effective diffusion coefficient and compare against the simulated dynamics.

The peptide sequence C(AGQ)₃W (denoted by CW₃) studied experimentally by Lapidus *et al.*¹ is simulated using all atom molecular dynamics in explicit solvent. A closely related work by Yeh and Hummer³⁶ was recently published that presents extensive all atom simulations of the shorter peptides C(AGQ)₁W and C(AGQ)₂W. The simulated mean quenching time for CW₃ is the same order of magnitude as found previously for the shorter peptides:³⁶ for diffusion limited quenching with a reaction radius of 4 Å we find mean contact time of approximately 6 ns (adjusted to solvent viscosity $\eta^{\text{wat}} = 1.0$ cP). The experimentally measured quenching time for this peptide (on the order of 100 ns) is many times longer than the diffusion limited contact time,¹ suggesting that the quenching rate for this system is reaction limited not diffusion limited. Indeed, the analysis of Yeh and Hummer³⁶ shows that the simulation is consistent with experiment assuming short distance quenching with a finite rate.³⁷ While this is very important for interpreting these particular experiments, the fundamental time scale that one wishes to ultimately understand in the folding context is the polymer dynamics of intrachain contacts, not the quenching properties of the probes. The simulation of CW₃ presented in this paper as well as those presented by Yeh and Hummer³⁶ show that a reduced effective diffusion coefficient is still required to fit even the shorter diffusion limited quenching time. This is the focus of the present paper as the reaction limited quenching is a rather separate issue.

The analysis of the present simulation suggests that the small effective diffusion coefficient does not originate from local internal friction or longitudinal ruggedness that can still be modeled in an otherwise Gaussian formalism. The accuracy of the single variable approximation to the multi-bead quenching time does not seem essential to the discrepancy between calculated and measured rates either. For this simple reaction, the *prima facie* reaction coordinate of the quenching would be expected to be the end-to-end distance $R(t) = |\mathbf{R}(t)|$. Consequently, $R(t)$ is also the relevant quantity requiring an accurate dynamical description. Due to Eq.(1), the pair correlation $\langle R(t)R(0) \rangle$ for a Gaussian model can be found once $\langle R^2 \rangle$ and $\phi(t)$ are specified. From the simulations, it is found that although $P_{\text{eq}}(R)$ is roughly Gaussian and $\phi(t)$ is well described by superimposed modes, $\langle R(t)R(0) \rangle$ relaxes much more slowly than the Gaussian expression.

This is independent of a specific polymer model, relying only on an assumption of isotropic Gaussian dynamics. The simulated $\langle R(t)R(0) \rangle$ can be approximately fit for long times with Gaussian dynamics by reducing the diffusion coefficient to a value that also brings the simulations into agreement with the bounds on the mean contact time provided by the WF and SSS approximations.

The most fruitful extensions to the Gaussian formalism would likely include an additional time scale or reaction coordinate that influences the end-to-end dynamics, rather than improving the description of $\phi(t)$ through more elaborate Gaussian chain models. Furthermore, since the contact rate is directly related to the auto-correlation $\langle R(t)R(0) \rangle$, it may be more convenient theoretically to focus on this correlation than the intrachain mean first passage time to improve polymer dynamics models.

Gaussian Chains

Although the conclusions (and the analysis of the simulations) do not rely on a specific Gaussian polymer dynamics model, introducing a rather general Gaussian chain at the outset makes the formalism more clear. Consider a chain consisting of N monomers with positions $\{\mathbf{r}_i\}$ and isotropic Gaussian connectivity

$$\beta U[\{\mathbf{r}\}] = \frac{3}{2b^2} \sum_{ij} \mathbf{r}_i \cdot \Gamma_{ij} \cdot \mathbf{r}_j, \quad (4)$$

where $\beta = 1/k_B T$ is the inverse temperature, and b is the root mean square distance between adjacent monomers along the chain. The static correlations of monomer positions are determined by the coefficients Γ_{ij} : $\langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle / b^2 = \Gamma_{ij}^{-1}$, where $\langle \dots \rangle$ denotes an equilibrium average over the equilibrium distribution $P_{\text{eq}}[\{\mathbf{r}\}] \sim \exp[-\beta U[\{\mathbf{r}\}]]$. (In this paper, spatial isotropy of matrices is implicit, e.g., Γ_{ij} is the direct product of an $N \times N$ matrix and the 3×3 identity matrix.)

The dynamics are assumed to be over-damped and Markovian with a spatially independent friction matrix γ_{ij} . The probability distribution $P[\{\mathbf{r}\}, t]$ evolves according to the Smoluchowski equation

$$\partial_t P[\{\mathbf{r}\}, t] = \mathcal{D}[\{\mathbf{r}\}] P[\{\mathbf{r}\}, t], \quad (5)$$

where the diffusion operator is defined as

$$\mathcal{D}[\{\mathbf{r}\}] = \sum_{ij} \frac{\partial}{\partial \mathbf{r}_i} P_{\text{eq}}[\{\mathbf{r}\}] \cdot D_{ij} \cdot \frac{\partial}{\partial \mathbf{r}_j} \frac{1}{P_{\text{eq}}[\{\mathbf{r}\}]} \quad (6)$$

with the diffusion matrix $D_{ij} = k_B T \gamma_{ij}^{-1}$.

For a Gaussian chain, the equilibrium correlation functions and averages for any functions $f(\mathbf{R})$ and $g(\mathbf{R})$ of the end-to-end vector $\mathbf{R} = \mathbf{r}_N - \mathbf{r}_i$ can be calculated as

$$\langle f(t)g(0) \rangle = \int d^3 \mathbf{R} \int d^3 \mathbf{R}_0 f(\mathbf{R}) g(\mathbf{R}_0) G(\mathbf{R}t|\mathbf{R}_0) P_{\text{eq}}(\mathbf{R}_0) \quad (7)$$

$$\langle f \rangle = \int d^3 \mathbf{R} f(\mathbf{R}) P_{\text{eq}}(\mathbf{R}), \quad (8)$$

where the Green's function $G(\mathbf{R}t|\mathbf{R}_0)$ and the equilibrium distribution $P_{\text{eq}}(\mathbf{R})$ are given by Eq.(1) and Eq.(3), respectively. Note that the chain dynamics are included through the dependence of $G(\mathbf{R}t|\mathbf{R}_0)$ on the pair correlation function $\phi(t)$. The form of $\phi(t)$ for this general Gaussian chain is given in the Appendix.

End-to-end quenching is incorporated into the Smoluchowski equation as a distance dependent sink term

$$\partial_t P[\{\mathbf{r}\}, t] = \mathcal{D}[\{\mathbf{r}\}] P[\{\mathbf{r}\}, t] - \epsilon k(\mathbf{R}) P[\{\mathbf{r}\}, t], \quad (9)$$

where $\epsilon k(\mathbf{R})$ is the quenching rate for the relative end-to-end separation \mathbf{R} . Following Pastor *et al.*,³³ I consider a Heavyside sink

$$k(\mathbf{R}) = \begin{cases} 1 & , \quad |\mathbf{R}| \leq a \\ 0 & , \quad |\mathbf{R}| > a \end{cases} \quad (10)$$

with diffusion limited quenching, $\epsilon \rightarrow \infty$. The reaction rate can be characterized by the mean first passage time

$$\tau = \int_0^\infty dt S(t) = \hat{S}(\omega = 0), \quad (11)$$

where $\hat{S}(\omega)$ is the Laplace transform of the survival probability

$$S(t) = \int d^3 \{\mathbf{r}\} P[\{\mathbf{r}\}, t], \quad (12)$$

$$\hat{S}(\omega) = \text{LT}_{t \rightarrow \omega}[S(t)] \equiv \int_0^\infty dt e^{-\omega t} S(t).$$

The Laplace transformed survival probability (and, hence, mean first passage time) for the Gaussian chain obey complementary variational bounds. To develop this relationship, it is necessary to introduce the adjoint Smoluchowski operator,

$$\mathcal{L}[\{\mathbf{r}\}] = - \sum_{ij} \frac{1}{P_{\text{eq}}[\{\mathbf{r}\}]} \frac{\partial}{\partial \mathbf{r}_i} P_{\text{eq}}[\{\mathbf{r}\}] \cdot D_{ij} \cdot \frac{\partial}{\partial \mathbf{r}_j}, \quad (13)$$

which is defined through $\mathcal{D}(P_{\text{eq}} f) = -P_{\text{eq}} \mathcal{L} f$, for any function $f[\{\mathbf{r}\}, t]$. Writing the probability distribution as $P(t) = \rho(t) P_{\text{eq}}$ and assuming equilibrium initial conditions, Eq.(9) becomes

$$[\omega + \mathcal{L}[\{\mathbf{r}\}] + \epsilon k(\mathbf{R})] \hat{\rho}[\{\mathbf{r}\}, \omega] = 1 \quad (14)$$

where $\hat{\rho} = \text{LT}_{t \rightarrow \omega}[\rho]$. In this notation the survival probability is $\hat{S}(\omega) = \langle \hat{\rho} \rangle$.

The variational bounds on the survival probability derived in Ref. 38 can be written in terms of trial functions $\varphi[\{\mathbf{r}\}]$ and $\xi[\{\mathbf{r}\}]$

$$F_\omega[\varphi] \leq \hat{S}(\omega) \leq K_\omega[\xi], \quad (15)$$

where the lower bound is given by

$$F_\omega[\varphi] = -\langle \varphi(\omega + \mathcal{L} + \epsilon k) \varphi \rangle + 2\langle \varphi \rangle, \quad (16)$$

and the upper bound is given by³⁹

$$\frac{1}{K_\omega[\xi]} = \omega + \frac{\epsilon \langle \xi k \rangle^2}{\langle \xi^2 k \rangle + \text{LT}_{t \rightarrow \omega} \epsilon \langle \Delta[\xi(t)k(t)] \Delta[\xi(0)k(0)] \rangle} \quad (17)$$

with $\Delta[\xi(t)k(t)] = \xi(t)k(t) - \langle \xi k \rangle$. The optimal trial functions $\delta F_\omega[\varphi^*] = 0$ and $\delta K_\omega[\xi^*] = 0$ satisfy Eq.(14), with $\varphi^* = \xi^* = \hat{\rho}$ and $F_\omega[\varphi^*] = K_\omega[\xi^*] = \hat{S}(\omega)$. Note that Eq.(15) evaluated at $\omega = 0$ are bounds on the mean first passage time $\tau = \hat{S}(0)$.

In practice, the bounds on $\hat{S}(\omega)$ and τ provided by Eqs.(15–17) are determined by the choice of trial function to optimize. The upper bound on τ was first derived by Doi⁴⁰ who noted that $K_{\omega=0}[\xi]$ evaluated at the simplest trial function $\xi = 1$ gives the well known Wilemski-Fixman approximation.³¹ For the diffusion limited reaction ($\epsilon \rightarrow \infty$), the WF approximation becomes

$$\tau_{\text{WF}} = \int_0^\infty dt \left(\frac{\langle k(t)k(0) \rangle}{\langle k \rangle^2} - 1 \right). \quad (18)$$

Similarly, the single variable approximation (like SSS) can be seen to be a lower bound on τ . Restricting the trial function to functions of the end-to-end relative vector, $F_\omega[\varphi(\mathbf{R})]$ can be reduced to averages over $P_{\text{eq}}(\mathbf{R})$ which are denoted by $\langle \dots \rangle_{\mathbf{R}}$. This is immediately evident for each term in Eq.(16) except $\langle \varphi \mathcal{L} \varphi \rangle$ which reduces according to

$$\begin{aligned} \langle \varphi \mathcal{L} \varphi \rangle &= \sum_{ij} \left\langle \frac{\partial \varphi(\mathbf{R})}{\partial \mathbf{r}_i} \cdot D_{ij} \cdot \frac{\partial \varphi(\mathbf{R})}{\partial \mathbf{r}_j} \right\rangle \\ &= \left\langle \frac{\partial \varphi(\mathbf{R})}{\partial \mathbf{R}} \cdot D_{\text{eff}} \cdot \frac{\partial \varphi(\mathbf{R})}{\partial \mathbf{R}} \right\rangle_{\mathbf{R}} \\ &= \langle \varphi \mathcal{L}(\mathbf{R}) \varphi \rangle_{\mathbf{R}}. \end{aligned} \quad (19)$$

Thus, for this class of trial function the lower bound becomes (for $\omega = 0$)

$$F_{\omega=0}[\varphi(\mathbf{R})] = -\langle \varphi(\mathcal{L}(\mathbf{R}) + \epsilon k) \varphi \rangle_{\mathbf{R}} + 2\langle \varphi \rangle_{\mathbf{R}}, \quad (20)$$

with the single variable adjoint operator

$$\mathcal{L}(\mathbf{R}) = -\frac{1}{P_{\text{eq}}(\mathbf{R})} \frac{\partial}{\partial \mathbf{R}} P_{\text{eq}}(\mathbf{R}) \cdot D_{\text{eff}} \cdot \frac{\partial}{\partial \mathbf{R}}, \quad (21)$$

and effective diffusion constant

$$D_{\text{eff}} = D_{NN} + D_{11} - D_{1N} - D_{N1}. \quad (22)$$

Optimizing Eq.(20) with respect to $\varphi(\mathbf{R})$ gives⁴¹ $[\mathcal{L}(\mathbf{R}) + \epsilon k] \varphi = 1$. For diffusion limited reactions ($\epsilon \rightarrow \infty$) and a Heavyside sink, the sink term can be represented by the boundary condition: $\varphi(\mathbf{R}) = 0$, for $|\mathbf{R}| \leq a$. Thus, the starting equation introduced by Szabo, Schulten and Schulten³²

$$\mathcal{L}(\mathbf{R}) \varphi = 1, \quad \varphi(|\mathbf{R}| \leq a) = 0, \quad (23)$$

is an approximation corresponding to a lower bound for the mean first passage time for intrachain contacts. The mean first passage time from Eq.(23) ($\tau = \langle \varphi \rangle$) can be solved exactly giving the SSS approximation³²

$$\tau_{\text{SSS}} = \frac{1}{D_{\text{eff}}} \int_a^\infty dx \frac{1}{x^2 \tilde{P}_{\text{eq}}(x)} \left[\int_x^\infty dy y^2 \tilde{P}_{\text{eq}}(y) \right]^2, \quad (24)$$

where $\tilde{P}_{\text{eq}}(R) \sim \exp[-\beta U(R)]$ is normalized over $a \leq R \leq \infty$, and D_{eff} is given by Eq.(22).

Putting this together, the mean first contact time for the original polymer is bounded by

$$\tau_{\text{SSS}} \leq \tau \leq \tau_{\text{WF}}. \quad (25)$$

Consider applying these bounds to a simulation of a peptide. Eq.(25) states that the simulated mean contact time is bounded by the SSS approximation calculated with the simulated potential of mean force and the WF approximation calculated from the simulated sink-sink correlation function. By itself, this is not very helpful in connecting to microscopic parameters of a polymer model that can be generalized to other problems, e.g., polymer based models of protein folding dynamics. However, the bounds calculated from the same many-bead polymer model are useful since the exact τ from the model is generally unknown. A chain with Gaussian dynamics is chosen in this paper to define the model and associated parameters that are compared with the simulated contact time and end-to-end dynamics.

In this case, the two approximations can be simplified further. For the SSS approximation, with $\beta U(R) = 3R^2/2\langle R^2 \rangle$, the inner integral of Eq.(24) can be evaluated giving³³

$$\tau_{\text{SSS}} D_{\text{eff}} = \frac{\langle R^2 \rangle}{6\Gamma[3/2, x_0]} \int_{x_0}^\infty du u^{-3/2} e^u \Gamma(3/2, u)^2, \quad (26)$$

where $\Gamma(a, x)$ is the incomplete gamma function, and

$$x_0 = \frac{3a^2}{2\langle R^2 \rangle}. \quad (27)$$

In the WF approximation, the sink-sink correlation function in Eq.(18) can be calculated using Eq.(7). For the case of a Heavyside sink, Wilemski and Fixman approximated the pair correlation using a “unbalanced Heavyside delta approximation”³¹ that can be written as

$$\frac{\langle k(t)k(0) \rangle}{\langle k \rangle^2} \approx \frac{\text{erf}(\xi) - (2/\sqrt{\pi})\xi \exp(-\xi^2)}{\text{erf}(\sqrt{x_0}) - (2/\sqrt{\pi})\sqrt{x_0} \exp(-x_0)}, \quad (28)$$

with x_0 given by Eq.(27), and

$$\xi(t) = \sqrt{\frac{x_0}{1 - \phi^2(t)}}. \quad (29)$$

The approximations Eq.(26) and Eqs.(28 - 29) were among those compared with simulations of Gaussian

chains by Pastor *et al.*;³³ this work verified (though did not emphasize) that the two approximations are bounds for the mean contact times.

Since the comparison of these approximations to measured (or simulated) contact times is framed in terms of the relative value of the diffusion coefficient, this parameter warrants closer consideration. As shown in the Appendix, a short time approximation to the diffusion coefficient also leads to D_{eff} given by Eq.(22). For a free draining chain (diagonal friction), $D_{\text{eff}} = D_N + D_1$ is the relative diffusion constant of the two end monomers, which becomes $D_{\text{eff}} = 2D_0$ if the friction is uniform ($D_{ij} = D_0\delta_{ij}$). Alternatively, Pastor *et al.*³³ also derived $D_{\text{eff}} = 2D_0$ for a Rouse chain (nearest neighbor connectivity) with uniform friction using a “local equilibrium” approximation, integrating over the internal monomer positions in Eq.(14); even though the formal manipulations are similar to the variational derivation, it is not evident from the “local equilibrium” approximation that the SSS approximation with this D_{eff} is actually a lower bound. The diffusion coefficient contained the WF approximation follows more naturally from the model through the time dependence of the pair correlation function $\phi(t)$.

The two approximations to τ given by Eq.(18) and Eq.(24) suggest (in slightly different ways) that the dynamics of the end-to-end distance $R(t) = |\mathbf{R}(t)|$ should be given careful attention. In the SSS approximation, Eq.(23) describes the dynamics of $R(t)$ in an isotropic potential of mean force $\beta U(R)$, i.e., $R(t)$ is treated as a single reaction coordinate. A single variable treatment should be fine to calculate rates as long as the effective diffusion constant is appropriate. In the formalism presented here, D_{eff} arises from reducing a more microscopic description to an effective one. On the other hand, without the advantage of a microscopic theory, D_{eff} could be determined from the dynamics of the reaction coordinate (obtained from a simulation, for example). Defined in this way, D_{eff} is truly an *effective* diffusion constant of the theory, not necessarily connected to microscopic parameters in a straight-forward way (see Ref. 42, for example). Turning now to the WF approximation, the dynamics of $R(t)$ enters explicitly through the sink-sink correlation function, since it is the autocorrelation of a function $R(t)$.

The pair correlation $\phi(t)$ has both a distance and angular component: $\phi(t) = \langle R(t)R(0) \cos \theta(t) \rangle / \langle R^2 \rangle$, where $\cos \theta(t)$ is the angle between $\mathbf{R}(t)$ and $\mathbf{R}(0)$. Separating the distance and angular components, we consider the correlation functions

$$\phi_d(t) = \frac{\langle R(t)R(0) \rangle}{\langle R^2 \rangle}, \quad \phi_\theta(t) = \langle \cos \theta(t) \rangle. \quad (30)$$

The Gaussian Green’s function [Eq.(1)] imposes a specific dependence of $\phi_d(t)$ and $\phi_\theta(t)$ on $\phi(t)$, just as it does on the sink-sink correlation function given in Eq.(28).

Performing the integrals in Eq.(7), gives

$$\phi_d^g(t) = \frac{2}{3\pi} \left[\frac{1 + 2\phi^2(t)}{\phi(t)} \arcsin \phi(t) + 3\sqrt{1 - \phi^2(t)} \right] \quad (31)$$

$$\phi_\theta^g(t) = \frac{2}{\pi} \left[\frac{2\phi^2(t) - 1}{\phi^2(t)} \arcsin \phi(t) + \frac{\sqrt{1 - \phi^2(t)}}{\phi(t)} \right] \quad (32)$$

(The superscript indicates that this is a Gaussian relationship.) Not only are each correlation a function of $\phi(t)$, but these Gaussian expressions obviously also have specific relationships to each other, e.g., $3\dot{\phi}_d^g(t) = \dot{\phi}(t)\phi_\theta^g(t)$.

Both $\phi_d^g(t)$ and $\phi_\theta^g(t)$ are normalized to one at $t = 0$. Expanding in powers of $\phi(t)$ gives to lowest order

$$\phi_d^g(t) \sim \frac{8}{3\pi} \left(1 + \frac{1}{6}\phi^2(t) \right), \quad \phi_\theta^g(t) \sim \frac{8}{3\pi}\phi(t). \quad (33)$$

Thus, the Gaussian correlation $\phi_d^g(t)$ approaches the long time limit $\langle R \rangle^2 / \langle R^2 \rangle = 8/3\pi$ at twice the rate that $\phi(t)$ decays, and $\phi_\theta^g(t)$ approaches zero at the same rate as $\phi(t)$.

Simulation

In this section, a molecular dynamics simulation of a short peptide is presented to investigate the end-to-end polymer dynamics and the diffusion limited quenching rate approximations developed in the previous section. Although simplified models can be very helpful in determining the essential underlying physics responsible for realistic polymer dynamics, the simulation reported here is of a detailed model of the CW₃ peptide studied experimentally by Lapidus *et al.*¹ As emphasized by Yeh and Hummer,³⁶ simulations of a realistic model provide a stringent test for the molecular potentials when compared directly with measured quenching rates. From a more theoretical point of view, simulations obviously also make accessible quantities essential to the consistency of the theory (e.g., correlation functions) that are not always easy to measure in the laboratory. Since this work is motivated primarily by the unexplained parameter values needed to fit experimentally measured rates, the simulation presented in this section is an attempt to capture the dynamics relevant to the experimental system.

The peptide Ace - C(AGQ)₃W - Nme is modeled by the force field of Cornell *et al.*⁴³ The AMBER 4.1⁴⁴ suite of programs are used, modified to include the generalized reaction field treatment of electrostatics^{45,46} with a cut-off of 9 Å. The system, solvated with approximately 1700 TIP3P water molecules, was held at constant temperature ($T = 300$ K) and pressure ($P = 1$ atm) by Berendsen couplings,⁴⁷ each with a relaxation time of 0.1 ps. The integration step is 0.002 ps. Non-bonded pair list was updated every 10 integration steps. A single trajectory of 50 ns was generated, starting from an extended conformation, with conformational coordinates saved every 0.5 ps.

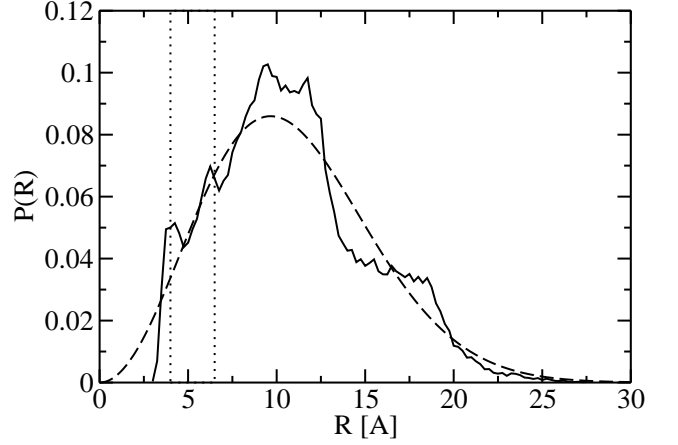


FIG. 1: Probability distribution of the end-to-end distance for CW₃. Solid: histogram from simulation with a bin width of 0.25 Å; dashed: Gaussian distribution with simulated mean square distance $\langle R^2 \rangle = 140 \text{ Å}^2$. The vertical dotted lines indicate the range of contact radii $4 \text{ Å} \leq a \leq 6.5 \text{ Å}$.

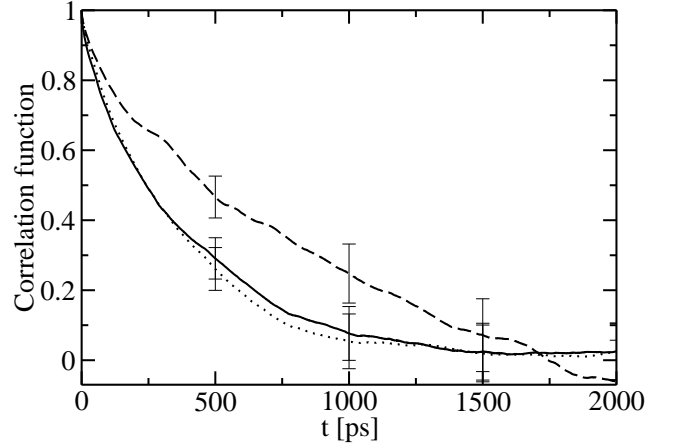


FIG. 2: Simulated pair correlation functions of CW₃. Solid: $\phi(t)$; dotted: $\phi_\theta(t)$; dashed: normalized $\bar{\phi}_d(t)$. Error bars are computed according to Zwanzig and Ailawadi⁴⁸; for example, the error bars for $\phi_\theta(t)$ are $\pm \sqrt{2T_\theta/T(1 - \phi_\theta(t))}$ where T is total simulation time (50 ns), and $T_\theta = \int dt \phi_\theta^2(t)$.

The geometric center of all heavy atoms of the cystine and tryptophan sidechains (denoted by cm) are chosen to define the end-to-end vector, $\mathbf{R}(t) = \mathbf{r}_{\text{cm}}^{\text{w}}(t) - \mathbf{r}_{\text{cm}}^{\text{c}}(t)$. Fig. 1 shows the probability distribution of the end-to-end distance $R = |\mathbf{R}|$, as well as a Gaussian distribution with the simulated mean square end-to-end distance, $\langle R^2 \rangle_{\text{sim}} \approx 140 \text{ Å}^2$. Although there is some deviation, the distribution can roughly be described as Gaussian. The mean distance $\langle R \rangle_{\text{sim}}^2 / \langle R^2 \rangle_{\text{sim}} \approx 0.87$ is within a few percent of the Gaussian relationship $\langle R \rangle^2 / \langle R^2 \rangle = 8/3\pi \approx 0.85$.

The simulated pair correlation functions, $\phi(t)$, $\phi_d(t)$, and $\phi_\theta(t)$ are shown in Fig. 2. (Note the radial correlation function has been normalized: $\bar{\phi}_d(t) = \langle \delta R(t) \delta R(0) \rangle / \langle \delta R^2 \rangle$, with $\delta R(t) = R(t) - \langle R \rangle$.) The pair

correlation function $\phi(t)$ follows the relaxation of the rotational correlation function $\phi_\theta(t)$ in agreement with the limiting Gaussian expression [Eq.(33)]. The radial autocorrelation decays more slowly than the others. Although a more careful evaluation is considered later, this slower relaxation can be seen to be a manifestation of non-Gaussian dynamics of the end-to-end relative vector $\mathbf{R}(t)$ [*c.f.*, Eq.(33)].

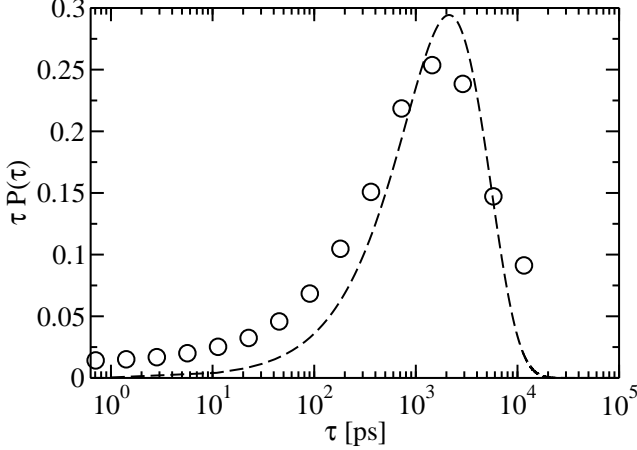


FIG. 3: First passage time distribution of CW_3 for a Heavy-side sink of radius $a = 5 \text{ \AA}$. The histogram is on a logarithmic scale with a bin width equal to powers of 2 (circles); Dashed line: exponential distribution, $P(\tau) = A\tau_{\text{sim}}^{-1} \exp[-\tau/\tau_{\text{sim}}]$, with $\tau_{\text{sim}} = \langle \Delta^2 \rangle_{\text{sim}} / 2\langle \Delta \rangle_{\text{sim}} = 2123 \text{ ps}$, and $A = 0.8$.

The diffusion limited mean first passage time can be calculated from the distribution of time intervals between contacts⁴⁹ $\{\Delta_k\}$:

$$\{\Delta_k\} = \{t_j - t_i | R(t_{i-1}) \leq a; R(t_i) > a; R(t_j) \leq a\}, \quad (34)$$

for $i < j$, and $k = 1 \dots n_{\text{int}}$. Taking the initial distribution to be every conformation with $R(t) > a$ gives the equilibrium distribution of contact times $P(\tau)$ with a mean contact time $\tau_{\text{sim}} = \langle \Delta^2 \rangle_{\text{sim}} / 2\langle \Delta \rangle_{\text{sim}}$. The simulated mean first passage times of CW_3 for different sink radii is given in Table I. For example, a sink radius of $a = 5.0 \text{ \AA}$ had $n_{\text{int}} = 870$ separate contact events with a mean contact time $\tau_{\text{sim}} = 2.1 \pm 0.7 \text{ ns}$. As shown in Fig. 3, the distribution of $P(\tau)$ for a sink radius of $a = 5.0 \text{ \AA}$ has a strong exponential component (seen by the width of a couple of decades) and a non-exponential component at short times. A non-exponential component is expected because of short-time recrossings at the sharp edge defining the Heavy-side sink. Although the two components are not well separated, the exponential component is described reasonably well by $P(\tau) = \tau_{\text{sim}}^{-1} \exp(-\tau/\tau_{\text{sim}})$ with an amplitude of 0.8. The simulated distribution $P(\tau)$ for other sink radii are similar. Note that the times reported are in the units from the simulation. Anticipating the discussion in the next section, to approximately convert this time to units consistent with the viscosity of water at 298 K ($\eta^{\text{wat}} = 1.0 \text{ cP}$), τ_{sim} should multi-

plied by a factor of two; for example, the mean time to make a contact at a distance $a = 4.0 \text{ \AA}$ corresponds to $\tau_{\text{sim}}^{\text{wat}} = 5.6 \pm 1.4 \text{ ns}$.

Free Diffusion Coefficient

A good estimate of the diffusion coefficient of the end monomers in TIP3P water is needed to compare the simulated contact times with analytical models. To this end, two separate 10 ns simulations of $H_2 - C - OH$ and $H_2 - W - OH$ in approximately 1180 TIP3P waters were carried out. Otherwise, the conditions are the same as those described for the peptide simulation. The value determined for the free diffusion coefficient D_0 appropriate for the simulation will be used in the next section to compare with analytic models.

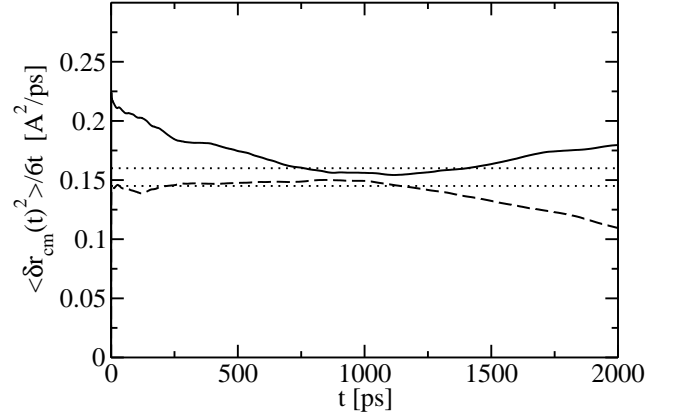


FIG. 4: Simulated diffusion coefficient for cystine (solid) and tryptophan (dashed) in TIP3P water. The dotted lines show the mean values over 500 ps to 1500 ps, $D_0^c = 0.16 \text{ \AA}^2/\text{ps}$ and $D_0^w = 0.145 \text{ \AA}^2/\text{ps}$.

The translational diffusion coefficient of the geometric center of the heavy atoms of the residues can be calculated from the Einstein relation $D \sim \langle \delta \mathbf{r}_{\text{cm}}(t)^2 \rangle / 6t$ for long times, where $\delta \mathbf{r}_{\text{cm}}(t) = \mathbf{r}_{\text{cm}}(t) - \mathbf{r}_{\text{cm}}(0)$. The diffusion coefficient for cystine and tryptophan are shown in Fig. 4. Although there is still some variation due to the finite time of the simulation, the average diffusion coefficient over 500ps to 1500ps is taken as an estimate, giving $D_0^c = 0.16 \text{ \AA}^2/\text{ps}$ for cystine and $D_0^w = 0.145 \text{ \AA}^2/\text{ps}$ for tryptophan. We can verify these values using Stokes Law, $k_B T / D_0 = 6\pi\eta_0 a_0$, with a_0 equal to the van der Waals radii of the sidechains,⁵⁰ to find the viscosity of TIP3P water, η_0 . For cystine, using $a_0 = 2.75 \text{ \AA}$ gives $\eta_0 = 0.50 \text{ cP}$; and for tryptophan, using $a_0 = 3.38 \text{ \AA}$ gives $\eta_0 = 0.45 \text{ cP}$; the two estimates differing by 10%. This is in reasonable agreement with the viscosity of TIP3P calculated by Shen and Freed, $\eta_0 = 0.506 \pm 0.043 \text{ cP}$.⁵¹

The effective diffusion constant for the single variable (SSS) approximation of the mean first contact time is taken to be the relative diffusion constant of cystine

and tryptophan, $D_{\text{eff}} = D_0^c + D_0^w = 3.0 \times 10^{-5} \text{ cm}^2/\text{s}$. Translating this to a uniform monomer friction we have $D_0 = D_{\text{eff}}/2 = 1.5 \times 10^{-5} \text{ cm}^2/\text{s}$. This value for the effective diffusion constant is the central assumption in the comparison of the simulation with Gaussian chain models.

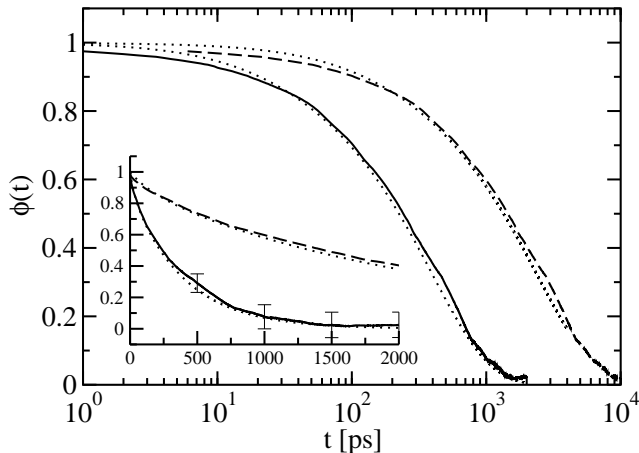


FIG. 5: Pair correlation $\phi(t)$ for CW_3 . Lower curves: $\phi(t)$ from simulation (solid), and Rouse chain (dotted) with $N = 11$, $b = 3.8 \text{ \AA}$, and $D_{ij} = D_0 \delta_{ij}$, $D_0 = 1.5 \times 10^{-5} \text{ cm}^2/\text{s}$. Upper curves, scaled diffusion coefficient, $\bar{D}_0 = D_0/6$: $\phi(t)$ from simulation with time scaled $t \rightarrow \bar{t} = (D_0/\bar{D}_0)t$ (dashed), and Rouse chain [Eqs.(43–45)] with monomer diffusion coefficient \bar{D}_0 (dotted). Inset is same plot on linear time axis.

Identifying D_{eff} as the relative diffusion constant of the two end monomers is only rigorously true for free-draining chains [Eq.(22)]. Comparison of the pair correlation function $\phi(t)$ from simulation and a free-draining chain supports for this choice (see Fig. 5). For times longer than about 50 ps, $\phi_{\text{sim}}(t)$ agrees closely with $\phi(t)$ from a Rouse chain [Eqs.(43–45)] with diagonal friction, $D_{ij} = D_0 \delta_{ij}$. The Rouse chain model uses reasonable parameters for this peptide, where the monomers of the chain are the α -carbons of the peptide: $N = 11$, $b = 3.8 \text{ \AA}$, and $D_0 = 1.5 \times 10^{-5} \text{ cm}^2/\text{s}$ is the monomer diffusion coefficient determined above. The mean square distance of this model Rouse chain, $\langle R^2 \rangle = (N-1)b^2 = 144 \text{ \AA}^2$, agrees with the simulated value of $\langle R^2 \rangle_{\text{sim}} = 140 \text{ \AA}^2$. The close agreement of the static and dynamic correlations is quite surprising, and one shouldn't expect it to hold generally, e.g., for all chain lengths. The correspondence with a Rouse chain is considered only to support the choice of the effective diffusion constant $D_{\text{eff}} = 2D_0$, with $D_0 = 1.5 \times 10^{-5} \text{ cm}^2/\text{s}$ — the remaining analysis assumes only this value for D_{eff} and Gaussian statistics of $\mathbf{R}(t)$.

Except where noted, the reported values for the diffusion coefficient and contact time correspond to the simulation viscosity $\eta_0 \approx 0.5 \text{ cP}$. To approximately convert to the viscosity of water at 298 K and 1 atm, $\eta^{\text{wat}} = 1.0 \text{ cP}$, the diffusion coefficient D_0 or contact time τ should be divided or multiplied by a factor of two, respectively. For

TABLE I: Diffusion limited mean first passage times of CW_3 with different sink radii.

$a[\text{\AA}]$	Simulation		$D_0 = 0.15 \text{ \AA}^2/\text{ps}$		$\bar{D}_0 = D_0/6$	
	n_{int}^a	$\tau_{\text{sim}} [\text{ns}]^b$	$\tau_{\text{SSS}} [\text{ns}]^c$	$\tau_{\text{WF}} [\text{ns}]^d$	$\tau_{\text{SSS}} [\text{ns}]^e$	$\tau_{\text{WF}} [\text{ns}]^f$
4.0	816	2.8 ± 0.7	0.27	0.67	1.6	4.0
4.5	901	2.5 ± 0.7	0.23	0.61	1.4	3.7
5.0	870	2.1 ± 0.7	0.20	0.56	1.2	3.4
5.5	1034	1.4 ± 0.3	0.17	0.51	1.0	3.1
6.0	1297	1.3 ± 0.3	0.15	0.46	0.90	2.8
6.5	1495	0.9 ± 0.2	0.13	0.42	0.78	2.5

^aNumber of contact events in simulation.

^bUncertainty is estimated as the standard error in the mean found by the limiting block averaged error using the contact times of each configuration with $R(t) \geq a$.

^cEqs.(26–27) using the simulated mean square distance, $\langle R^2 \rangle_{\text{sim}} = 140 \text{ \AA}^2$, and $D_{\text{eff}} = 2D_0$.

^dEq.(18) and Eqs.(28–29) using the simulated mean square distance, $\langle R^2 \rangle_{\text{sim}} = 140 \text{ \AA}^2$, and pair correlation, $\phi_{\text{sim}}(t)$.

^eSame as [c], but with $D_{\text{eff}} = 2\bar{D}_0$.

^fSame as [d], but with time scaled as $t \rightarrow \bar{t} = (D_0/\bar{D}_0)t$, i.e., using $\phi_{\text{sim}}(\bar{t})$.

example, the relative diffusion coefficient corresponding to $\eta^{\text{wat}} = 1.0 \text{ cP}$ is $2D_0^{\text{wat}} = 1.5 \times 10^{-5} \text{ cm}^2/\text{s}$.

Comparison of Simulation with Gaussian Dynamics

In this section, the simulated mean contact time and correlation functions of CW_3 are compared with Gaussian polymer models. Here, it is demonstrated that the effective diffusion coefficient required to approximately describe the radial autocorrelation function brings the bounding SSS and WF approximations into agreement with the simulated mean first contact time.

The parameters required to calculate the WF and SSS approximations can be chosen in a couple of different ways. For the lower bound, $\tau_{\text{SSS}} D_{\text{eff}}$ calculated with either the simulated potential of mean force $U(R) = -kT \log P_{\text{eq}}(R)$ [Eq.(24)], or the harmonic potential $U(R) = 3R^2/2\langle R^2 \rangle$ [Eqs.(26–27)] with $\langle R^2 \rangle$ from simulation or the Rouse chain agree to within 10% of each other for all the sink radii considered. This suggests that longitudinal ruggedness in the potential of mean force does not significantly reduce the effective diffusion coefficient for this peptide;⁵² this conclusion is also consistent with the analysis of quenching times for the two shorter peptides presented by Yeh and Hummer.³⁶ For the upper bound [Eq.(18) with the Gaussian correlations Eqs.(28–29)], τ_{WF} calculated with either $\phi(t)$ and $\langle R^2 \rangle$ from simulation or from the Rouse model differ by a few percent. Although these calculated rates are in sufficient agreement for the subsequent analysis, to emphasize independence of a Gaussian chain model, the reported mean first contact times use the parameters directly from the simulation: harmonic potential defined with $\langle R^2 \rangle_{\text{sim}}$ for τ_{SSS} , and $\phi_{\text{sim}}(t)$ and $\langle R^2 \rangle_{\text{sim}}$ for τ_{WF} .

As shown in Table I, the calculated SSS and WF

approximations underestimate the simulated mean first contact times for all sink radii considered. The SSS approximation is calculated with $D_{\text{eff}} = 2D_0$. If D_0 is treated as a fitting parameter (similar to previous analysis of experimentally measured times), the fit to τ_{SSS} defines a new effective monomer diffusion constant $\bar{D}_0 = D_0/\alpha_{\text{SSS}}$ with $\alpha_{\text{SSS}} = \tau_{\text{sim}}/\tau_{\text{SSS}}$. From Table I, α_{SSS} ranges from approximately 7 for $a = 6.5 \text{ \AA}$ to 10 for $a = 4.0 \text{ \AA}$. Following the same reasoning, the WF approximation also has an associated effective diffusion coefficient. For overdamped motion, an effective diffusion constant scales time as $\bar{t}\bar{D}_0 = tD_0$, so that using D_0 as a fitting parameter gives $\bar{D}_0 = D_0/\alpha_{\text{WF}}$ with $\alpha_{\text{WF}} = \tau_{\text{sim}}/\tau_{\text{WF}}$. From Table I, α_{WF} ranges from 2 for $a = 6.5 \text{ \AA}$ to 4 for $a = 4.0 \text{ \AA}$. (Note that $\alpha_{\text{SSS}} \geq \alpha_{\text{WF}}$ follows from $\tau_{\text{SSS}} \leq \tau_{\text{WF}}$.) Since SSS and WF are upper and lower bound approximations, respectively, the effective diffusion coefficient for the original chain should actually lie between the values obtained from each approximation; then, $\tau_{\text{SSS}} \leq \tau_{\text{sim}} \leq \tau_{\text{WF}}$. For example, for the sink radius $a = 5.0 \text{ \AA}$, the \bar{D}_0 should satisfy $D_0/10 \lesssim \bar{D}_0 \lesssim D_0/4$.

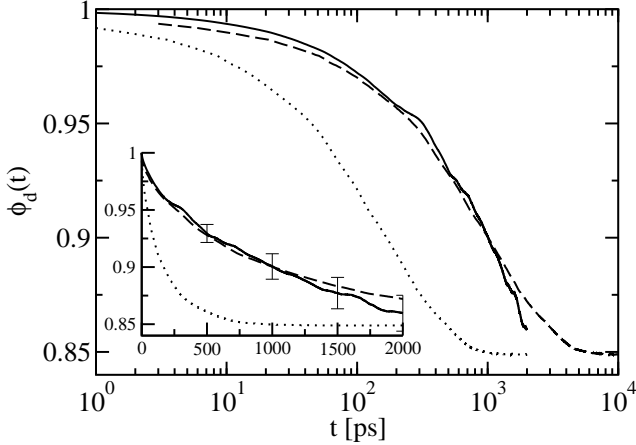


FIG. 6: Distance pair correlation $\phi_d(t)$ for CW3. Solid: simulation; dotted: $\phi_d^g(t)$ using the simulated pair correlation $\phi_{\text{sim}}(t)$ in Eq.(31); dashed: same as dotted curve with time scaled $t \rightarrow \bar{t} = (D_0/\bar{D}_0)t$ with $D_0/\bar{D}_0 = 6$. Inset is same plot on linear time axis.

Why do these approximations require a reduced effective diffusion coefficient to fit simulated contact rates? It is argued here that the effective diffusion coefficient reflects the slow dynamics of the end-to-end distance. In terms of comparing to a Gaussian model, the small effective diffusion coefficient accounts for the discrepancy of using a Gaussian theory to describe the non-Gaussian dynamics of the end-to-end vector. The simulated correlations $\phi_d(t)$ and $\phi_\theta(t)$ are shown as solid lines in Fig. 6 and Fig. 7, respectively. The dotted lines in these figures show the Gaussian expressions $\phi_d^g(t)$ and $\phi_\theta^g(t)$ [Eqs.(31–32)] calculated from the simulated pair correlations $\phi_{\text{sim}}(t)$. Comparing the simulated correlations with the Gaussian correlations tests the Gaussian statistics of $\mathbf{R}(t)$, [i.e., this checks whether correlations can be accurately cal-

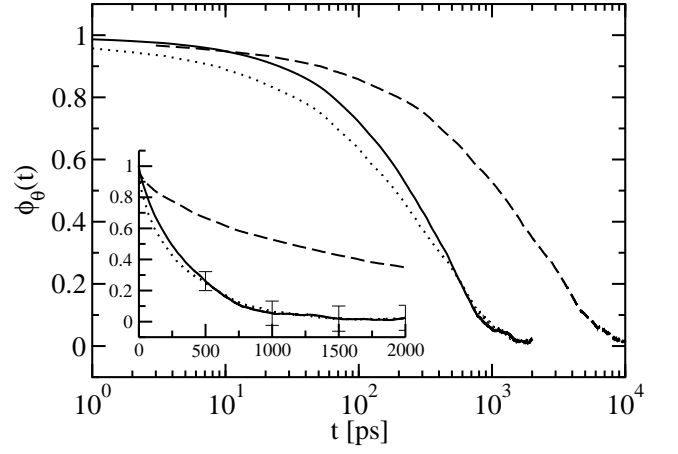


FIG. 7: Angular pair correlation $\phi_\theta(t)$ for CW3. Solid: simulation; dotted: $\phi_\theta^g(t)$ using the simulated pair correlation $\phi_{\text{sim}}(t)$ in Eq.(32); dashed: same as dotted curve with time scaled $t \rightarrow \bar{t} = (D_0/\bar{D}_0)t$ with $D_0/\bar{D}_0 = 6$. Inset is same plot on linear time axis.

culated with the isotropic Gaussian Green's function in Eq.(1)]. Both the simulated radial and angular correlations decay more slowly than the Gaussian model predicts. However, the deviation of the angular correlation from Gaussian form is small in comparison to the deviation of the radial correlations. The semi-log plot in Fig. 6 shows that the simulated $\phi_d(t)$ and $\phi_d^g(t)$ have roughly the same shape, but are translated along the time axis, i.e., they differ by a constant scaling of time. As shown by the dashed curve in Fig. 6, $\phi_d^g(t)$ agrees reasonably well with the simulated radial correlation if time scaled as $t \rightarrow \bar{t} = \alpha t$ with $\alpha = 6$. For overdamped motion, this is equivalent to an effective monomer diffusion coefficient $\bar{D}_0 = D_0/\alpha$. As shown in Fig. 5 and Fig. 7, this scaling applied to $\phi(t)$ and $\phi_\theta^g(t)$ results in a poor agreement with the simulated correlations. So, rescaling the diffusion constant to bring the Gaussian and simulated radial correlation function into agreement, makes $\phi(t)$ and $\phi_\theta^g(t)$ disagree with the simulated dynamics.

The dynamics of the end-to-end distance is the most relevant to mean first contact times calculations. Table I also shows the calculated times using the effective monomer diffusion coefficient⁵³ ($\bar{D}_0 = D_0/6$ with $\alpha = 6$) determined from the fit to $\phi_d(t)$. With this scaling, the SSS and WF approximations bracket the simulated contact time, as they should since they are upper and lower bounds. Since the approximations provide bounds for a rather broad range of α , these results offer only speculative support that the effective diffusion coefficient determined by the mean first contact time reflects the dynamics of $R(t)$. Because attention to $\phi_d(t)$ arises quite naturally from general considerations, the consistency of the analysis is very suggestive.

Conclusion and Outlook

The simulated diffusion limited mean first contact time agrees with the bounds provided by the WF and SSS approximations for a Gaussian chain if the diffusion coefficient is reduced. When viewed as completely independent approximations, it would seem that the SSS and WF approximations would each require different rescalings. It was shown, however, that the two approximations are complementary variational bounds, giving a range of values for the scaling because the approximations should merely bracket the measured times, not necessarily fit them each independently.

It was also demonstrated that the relative end-to-end distance of the simulated peptide relaxes more slowly than the relaxation from Gaussian dynamics. It is argued that the effective diffusion coefficient appropriate for the calculated mean first contact time also describes the end-to-end distance dynamics. Evidence supporting this connection is weakened by the broad interval between the bounds provided by the SSS and WF approximations. Tighter bounds (and other approximations) can be generated by the variational formalism through more flexible trial functions, helping to make support for this relationship more precise. If this relationship is indeed valid, one can focus on $\phi_d(t)$ rather than the mean first contact time itself to improve the model. Furthermore, the analysis demonstrates that accurately modeling $\phi(t)$ is not the complete description of the polymer dynamics, for the Gaussian statistics are checked using $\phi_{\text{sim}}(t)$ rather than the pair correlation from a polymer model.

It should be noted that Gaussian dynamics and the reduction of a multi-bead model to a single variable model are separate issues. The variational bounds $\tau_{\text{SSS}} \leq \tau \leq \tau_{\text{WF}}$ is a relationship between the mean contact time of a single variable model (with a specific D_{eff}) to the corresponding multi-bead model. Reducing the description to a single variable model provides the microscopic interpretation of D_{eff} as the relative free diffusion coefficient of the end monomers. In general, both models may be non-Gaussian, with the potential in the reduced model replaced by the potential of mean force of the chain. It seems that a microscopic interpretation of the slow end-to-end distance relaxation would require a multi-bead description, perhaps leading to an effective single variable description which is easier to handle analytically. In any case, the comparison of the end-to-end distance and angular relaxation should guide the development of a microscopic model that extends beyond the simple Gaussian dynamics of a single reaction coordinate.

In a recent paper, Lapidus *et al.*⁵⁴ present further experimental work and theoretical modeling of quenching in CW_k peptides. In these new experiments, the viscosity dependence of the quenching rate in CW_k peptides is measured, isolating the diffusion limited component of the measured quenching rate. Fitting the measured diffusion limited quenching rate with a single variable approximation corresponding to a stiff chain gives $\bar{D}_0 \approx D_0/10$.

This is the same factor obtained from the fit to the simulations presented here with a reaction radius of $a \leq 5 \text{ \AA}$ had the SSS approximation been used to define the effective monomer diffusion coefficient \bar{D}_0 (see Table I). To avoid confusion, it is emphasized again that the value found from our analysis of the simulation, $\bar{D}_0 = D_0/6$, was determined by the dynamics of the end-to-end distance, not the lower bound approximation to the contact time.

In the theoretical modeling portion of Lapidus *et al.*,⁵⁴ the mean first contact time of the end-to-end monomers is studied through Langevin simulations of a model that includes excluded volume and chain stiffness (provided by the dihedral angle potential), but no other non-local interactions between the monomers. The dynamics of the squared end-to-end distance from the simulation is analyzed in some detail with respect to the single reaction coordinate description. While the relationship between end-to-end distance dynamics and other correlation functions (angular correlations, for example) is not considered, Lapidus *et al.*⁵⁴ do emphasize the connection between the effective diffusion coefficient appropriate for the mean quenching time and the relaxation the end-to-end distance. Comparing to experimental data, their analysis suggests that the effective diffusion constant determined by a fit to the single variable approximation is not explained by the local properties of the backbone, but must involve interactions left out of the Langevin simulations; e.g., non-local contacts or explicit solvent. This very interesting conclusion limits the possible mechanism responsible for the slow non-Gaussian relaxation of the end-to-end distance.

The elements missing from these Langevin simulations are included in the all-atom molecular dynamics simulations analyzed in the present paper. Here, an improved model with gated diffusion seems promising: the trajectory of $R(t)$ shows relatively long waiting times separating periods of larger diffusional flights in distance reminiscent of continuous time random walks, for example. To be consistent with the simulated dynamics, such a model should have the property that the gating retards the end-to-end distance relaxation, but has a relatively smaller effect on the angular relaxation, at least for chains not too long compared to the persistence length. More refined analysis of the conformational transitions in the simulation should give insight into how to formulate a gated diffusion model; this is left for future work. Any microscopic model of the gated diffusion should be constrained by the reasonable expectation that the dynamics become Gaussian in the limit of very long chains.

Acknowledgments

I thank William Eaton making Ref. 54 available before publication. I also thank Angel Garcia for guidance in carrying out the simulations and helpful suggestions; and Peter Wolynes for our many discussions on this topic

and his insightful comments on the manuscript. This work was supported by a Director's Postdoctoral Fellowship and the Los Alamos National Laboratory Directed Research Fund (Ben McMahon, PI).

Appendix: $\phi(t)$ and the short time D_{eff} for Gaussian chains

This appendix gives the pair correlations $\phi(t)$ and derives the single variable approximation effective diffusion coefficient D_{eff} for a general Gaussian chain. The effective diffusion coefficient, derived here as a short time approximation, can also be obtained from the variational lower bound [Eq.(22)] as well as a “local equilibrium” approximation.³³

First, the diffusion matrix in the Smoluchowski equation [Eq.(5)] is written as a product of a diffusion constant and a unitless matrix $D_{ij} = D_0 \tilde{D}_{ij}$. The Smoluchowski equation can be diagonalized by transforming into normal modes $\mathbf{r}_i = \sum_p Q_{ip} \tilde{\mathbf{r}}_p$ where the transformation matrix \mathbf{Q} is defined by

$$\begin{aligned} \mathbf{Q}^{-1} \tilde{\mathbf{D}} \mathbf{r} \mathbf{Q} &= \text{diag}(\lambda_p) \\ \mathbf{Q}^{-1} \tilde{\mathbf{D}} \mathbf{Q}^{-1T} &= \text{diag}(\nu_p) \\ \mathbf{Q}^T \mathbf{r} \mathbf{Q} &= \text{diag}(\mu_p) \end{aligned} \quad (35)$$

with $\lambda_p = \nu_p \mu_p$, $p = (0 \dots N-1)$. The pair correlations of the end to end vector, $\mathbf{R} = \mathbf{r}_N - \mathbf{r}_1$, can then be written in terms of the normal modes

$$\phi(t) = \frac{b^2}{\langle R^2 \rangle} \sum_{p \neq 0} \frac{1}{\mu_p} (Q_{Np} - Q_{1p})^2 \exp \left[-\frac{3D_0}{b^2} \lambda_p t \right]. \quad (36)$$

Next, consider the effective diffusion constant of the single variable approximation to the Gaussian chain. The single variable approximation (such as SSS) treats $\mathbf{R}(t)$ as the only dynamical variable. The Green's function for $\mathbf{R}(t)$ [Eq.(1)] satisfies the diffusion equation

$$\partial_t G(\mathbf{R}t | \mathbf{R}_0) = \mathcal{D}(\mathbf{R}; t) G(\mathbf{R}t | \mathbf{R}_0), \quad (37)$$

where \mathcal{D} is the single variable diffusion operator

$$\mathcal{D}(\mathbf{R}; t) = D(t) \frac{\partial}{\partial \mathbf{R}} P_{\text{eq}}(\mathbf{R}) \cdot \frac{\partial}{\partial \mathbf{R}} \frac{1}{P_{\text{eq}}(\mathbf{R})} \quad (38)$$

with a time dependent diffusion coefficient

$$D(t) = -\frac{\langle R^2 \rangle}{3} \frac{d}{dt} \log \phi(t). \quad (39)$$

This equation is exact. The single variable approximation requires a choice for an effective diffusion constant D_{eff} to replace $D(t)$. Bicout and Szabo⁵⁵ discuss the choice of D_{eff} and demonstrate some limitations of the single variable approximation in the context of rate calculations of electron transfer in non-Debye solvents.

The expression for D_{eff} derived from the variational lower bound [Eq.(22)] is equivalent to a short time approximation, $D_{\text{eff}} = D(0)$. Using $\lambda_p = \nu_p \mu_p$ and $\tilde{D}_{ij} = \sum_p \nu_p Q_{ip} Q_{jp}$ from Eq.(35), $\dot{\phi}(0)$ reduces to

$$\dot{\phi}(0) = -\frac{3D_0}{\langle R^2 \rangle} \sum_{p \neq 0} \frac{\lambda_p}{\mu_p} (Q_{Np} - Q_{1p})^2 \quad (40)$$

$$= -\frac{3D_0}{\langle R^2 \rangle} (\tilde{D}_{NN} + \tilde{D}_{11} - \tilde{D}_{1N} - \tilde{D}_{N1}). \quad (41)$$

Since $\phi(0) = 1$, $D_{\text{eff}} = D(0)$ becomes

$$D_{\text{eff}} = D_{NN} + D_{11} - D_{1N} - D_{N1} \quad (42)$$

which is Eq.(22).

As a specific example, the Rouse chain (nearest neighbor connectivity) and uniform diagonal friction, $D_{ij} = D_0 \delta_{ij}$ has modes and relaxation rates

$$\begin{aligned} Q_{ip} &= \sqrt{\frac{2 - \delta_{p0}}{N}} \cos \left(\frac{p(i - 1/2)}{\pi N} \right) \\ \lambda_p &= \mu_p = 4 \sin^2 \left(\frac{p\pi}{2N} \right), \end{aligned} \quad (43)$$

and the mean square end-to-end distance is

$$\langle R^2 \rangle = (N-1)b^2. \quad (44)$$

The effective short time diffusion constant,

$$D_{\text{eff}} = 2D_0, \quad (45)$$

is the relative diffusion constant of the chain ends. Internal friction (of the form in Ref. 22) reduces D_{eff} .

In the single variable approximation to the end-to-end chain dynamics, the pair correlation decays with a single time constant

$$\phi(t) = \exp \left[-\frac{3D_{\text{eff}}}{\langle R^2 \rangle} t \right]. \quad (46)$$

¹ L. J. Lapidus, W. A. Eaton, and J. Hofrichter, Proc. Natl. Acad. Sci. USA **97**, 7220 (1999).

² O. Bieri, J. Wirz, B. Hellrung, M. Schutkowski, M. Drewello, and T. Kiefhaber, Proc. Natl. Acad. Sci. USA **96**, 9597 (1999).

³ R. R. Hudgins, F. Huang, G. Gramlich, and W. M. Nau, J. Am. Chem. Soc. **124**, 556 (2002).

⁴ E. Haas, E. Katchalski-Katzir, and I. Z. Steinberg, Biopolymers **17**, 11 (1978).

⁵ E. Haas, IEEE J. Quantum Electron. **2**, 1088 (1996).

- ⁶ M. I. Wallace, L. Ying, S. Balasubramanian, and D. Klennerman, *Proc. Natl. Acad. Sci. USA* **98**, 5584 (2001).
- ⁷ J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, *Proteins Struct. Funct. Genet.* **21**, 167 (1995).
- ⁸ S. E. Jackson, *Fold. Des.* **3**, R81 (1998).
- ⁹ M. Gruebele, *Annu. Rev. Phys. Chem.* **50**, 485 (1999).
- ¹⁰ L. S. Itzhaki, D. E. Otzen, and A. R. Fersht, *J. Mol. Biol.* **254**, 260 (1995).
- ¹¹ J. J. Portman, S. Takada, and P. G. Wolynes, *Phys. Rev. Lett.* **81**, 5237 (1998).
- ¹² O. V. Galzitskaya and A. V. Finkelstein, *Proc. Natl. Acad. Sci. USA* **96**, 112999 (1999).
- ¹³ E. Alm and D. Baker, *Proc. Natl. Acad. Sci. USA* **96**, 11305 (1999).
- ¹⁴ V. Muñoz and W. A. Eaton, *Proc. Natl. Acad. Sci. USA* **96**, 11311 (1999).
- ¹⁵ B. A. Shoemaker, J. Wang, and P. G. Wolynes, *J. Mol. Biol.* **287**, 675 (1999).
- ¹⁶ D. A. Debe and W. A. Goddard III, *J. Mol. Biol.* **294**, 619 (1999).
- ¹⁷ J. J. Portman, S. Takada, and P. G. Wolynes, *J. Chem. Phys.* **114**, 5082 (2001).
- ¹⁸ V. Muñoz, *Curr. Opin. Struct. Biol.* **11**, 212 (2001).
- ¹⁹ D. E. Makarov, C. A. Keller, K. W. Plaxco, and H. Metiu, *Proc. Natl. Acad. Sci. USA* **99**, 3535 (2002).
- ²⁰ R. Zwanzig, *J. Chem. Phys.* **60**, 2717 (1974).
- ²¹ M. Bixon and R. Zwanzig, *J. Chem. Phys.* **68**, 1896 (1978).
- ²² C. W. Manke and M. C. Williams, *Macromolecules* **18**, 2045 (1985).
- ²³ M. Fixman, *J. Chem. Phys.* **89**, 2442 (1988).
- ²⁴ R. Zwanzig, *Proc. Natl. Acad. Sci. USA* **85**, 2029 (1988).
- ²⁵ J. D. Bryngelson and P. G. Wolynes, *J. Phys. Chem.* **93**, 6902 (1989).
- ²⁶ D. Thirumalai, V. Ashwin, and J. K. Bhattacharjee, *Phys. Rev. Lett.* **77**, 5385 (1996).
- ²⁷ S. Takada, J. J. Portman, and P. G. Wolynes, *Proc. Natl. Acad. Sci. USA* **94**, 2318 (1997).
- ²⁸ T. R. Kirkpatrick and D. Thirumalai, *Phys. Rev. B* **36**, 5388 (1987).
- ²⁹ Very recently,³⁰ this problem has been reduced to the solution of an integral equation for the special case when the quenching has a delta function distance dependence.
- ³⁰ I. M. Sokolov, *cond-mat/0207159* (2002).
- ³¹ G. Wilemski and M. Fixman, *J. Chem. Phys.* **60**, 878 (1973).
- ³² A. Szabo, K. Schulten, and Z. Schulten, *J. Chem. Phys.* **72**, 4350 (1980).
- ³³ R. W. Pastor, R. Zwanzig, and A. Szabo, *J. Chem. Phys.* **105**, 3878 (1996).
- ³⁴ G. Srinivas, A. Yethiraj, and B. Bagchi, *J. Chem. Phys.* **114**, 9170 (2001).
- ³⁵ A. V. Barzykin, K. Seki, and M. Tachiya, *J. Chem. Phys.* **117**, 1377 (2002).
- ³⁶ I.-C. Yeh and G. Hummer, *J. Am. Chem. Soc.* **124**, 6563 (2002).
- ³⁷ L. J. Lapidus, W. A. Eaton, and J. Hofrichter, *Phys. Rev. Lett.* **87**, 258101 (2001).
- ³⁸ J. J. Portman and P. G. Wolynes, *J. Phys. Chem. A* **103**, 10602 (1999).
- ³⁹ This form of the upper bound is given in Ref. 38 Eq.(B9) with $g = \xi$.
- ⁴⁰ M. Doi, *Chem. Phys.* **11**, 107 (1975).
- ⁴¹ The variational bounds can be similarly developed for the single variable equation $[\mathcal{L}(\mathbf{R}) + \epsilon k] \hat{\rho} = 1$. For finite ϵ , bounds of this one-dimensional problem generated from different trial functions has been addressed in Ref. 38. We do not pursue this approach for Eq.(23) because the diffusion limited rate can be calculated exactly through Eq.(24).
- ⁴² N. D. Socci, J. N. Onuchic, and P. G. Wolynes, *J. Chem. Phys.* **104**, 5860 (1996).
- ⁴³ W. D. Cornell, P. Cieplak, C. I. Bailey, I. R. Gould, K. M. Mertz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. Kollman, *J. Am. Chem. Soc.* **117**, 5179 (1995).
- ⁴⁴ D. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, I. T. E. Cheatham, D. M. Ferguson, U. C. Singh, P. Weiner, and P. Kollman, *AMBER 4.1* (1995).
- ⁴⁵ L. R. Pratt, G. Hummer, and A. E. Garcia, *Biop. Chem.* **51**, 147 (1994).
- ⁴⁶ G. Hummer, D. M. Soumpasis, and M. Neuman, *J. Phys. Cond. Matt.* **6**, A141 (1994).
- ⁴⁷ H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *J. Chem. Phys.* **81**, 3684 (1984).
- ⁴⁸ R. Zwanzig and N. K. Ailawadi, *Phys. Rev.* **182**, 280 (1969).
- ⁴⁹ L. J. Lapidus, W. A. Eaton, and J. Hofrichter, *J. Mol. Biol.* **319**, 19 (2002).
- ⁵⁰ T. E. Creighton, *Proteins: Structures and Molecular Properties* (W. H. Freeman and Company, NY, 1993), 2nd ed.
- ⁵¹ M. Shen and K. F. Freed, *Biophys. J.* **82**, 1791 (2002).
- ⁵² The agreement, however, is to some degree accidental since the SSS approximation depends in part on the weight of $P_{eq}(R)$ at small distances, where the Gaussian is a poor approximation to the simulated potential of mean force. The additional weight of the Gaussian at small R is offset to some extent by the peak centered around 4 Å in the simulated $P_{eq}(R)$.
- ⁵³ Converting the effective diffusion coefficient to the viscosity $\eta^{\text{wat}} = 1.0$ cP, gives $2D_0^{\text{wat}} = 1.5 \times 10^{-5} \text{ cm}^2/\text{s}$ and $2\bar{D}_0^{\text{wat}} = 2.5 \times 10^{-6} \text{ cm}^2/\text{s}$.
- ⁵⁴ L. J. Lapidus, P. J. Steinbach, W. Eaton, A. Szabo, and J. Hofrichter, *J. Phys. Chem. B* (2002), submitted.
- ⁵⁵ D. J. Bicout and A. Szabo, *J. Chem. Phys.* **109**, 2325 (1998).